Supplementary Appendix to:

Learning can generate Long Memory

Guillaume Chevillon* Sophocles Mavroeidis†

ESSEC Business School & CREST University of Oxford

December 1, 2016

Contents

1 WLS interpretation of DGLS **2**

2 Proof of Expression (11) **3**

3 Degree of Long Memory of y_t^2 **4**

4 Additional Simulations **12**

*ESSEC Business School, Department of Information Systems, Decision Sciences and Statistics, Avenue Bernard Hirsch, BP50105, 95021 Cergy-Pontoise cedex, France. Email: chevillon@essec.edu.

†Department of Economics and INET at Oxford, University of Oxford, Manor Road, Oxford, OX1 3UQ, United Kingdom. Email: sophocles.mavroeidis@gmail.com.

1 WLS interpretation of DGLS

The model of Malmendier and Nagel (2016) assumes a non-increasing gain algorithm with gain sequence $g_t = \min(1, \theta/t)$ where θ denotes the forgetting factor. So, denoting by $\lceil\theta\rceil$ the *ceiling* of θ (i.e., the smallest integer at least as large as θ), expression (6) implies

$$y_{t+1}^e = \frac{\theta}{\lceil\theta\rceil} y_{\lceil\theta\rceil} \prod_{i=0}^{t-\lceil\theta\rceil-1} (1-g_{t-i}) + \sum_{j=0}^{t-\lceil\theta\rceil-1} \left[g_{t-j} \prod_{i=0}^{j-1} (1-g_{t-i}) \right] y_{t-j}.$$

Hence y_{t+1}^e can be written as

$$y_{t+1}^e = \sum_{j=0}^{t-\lceil\theta\rceil} \kappa_{t,j} y_{t-j},$$

where

$$\kappa_{t,j} = \begin{cases} \frac{\theta}{t-j} \prod_{i=t-j+1}^{t} \frac{i-\theta}{i}, & \text{if } j < t - \lceil\theta\rceil; \\ \frac{\theta}{\lceil\theta\rceil} \prod_{i=\lceil\theta\rceil+1}^{t} \frac{i-\theta}{i} & \text{if } j = t - \lceil\theta\rceil; \\ 0, & \text{if } j > t - \lceil\theta\rceil. \end{cases}$$

Since $q(q+1)\ldots(q+n) = \frac{\Gamma(q+n+1)}{\Gamma(q)}$ if q is not a negative integer, the weights satisfy, for $t > \lceil\theta\rceil$,

$$\kappa_{t,j} = \begin{cases} \theta \frac{\Gamma(t+1-\theta)}{\Gamma(t-j+1-\theta)} \frac{\Gamma(t-j)}{\Gamma(t+1)}, & \text{if } j \leq t - \lceil\theta\rceil; \\ 0, & \text{if } j > t - \lceil\theta\rceil. \end{cases}$$

We now ask whether there exists a sequence $w_{t,j}$ such that y_{t+1}^e can be written as solution to the Weighted Least-Squares problem:

$$y_{t+1}^e = \underset{a}{\operatorname{argmin}} \sum_{j=0}^{t-1} w_{t,j} (y_{t-j} - a)^2, \quad \sum_{j=0}^{t-1} w_{t,j} = 1.$$

i.e. such that

$$y_{t+1}^e = \sum_{j=0}^{t-1} w_{t,j} y_{t-j},, \quad \sum_{j=0}^{t-1} w_{t,j} = 1.$$

By identification, it is clear that it suffices to choose

$$w_{t,j} = \frac{\kappa_{t,j}}{\sum_{j=0}^{t-1} \kappa_{t,j}}. \tag{S-1}$$

We now consider the asymptotic behavior of the latter weights. Using Stirling's formula that states that $\frac{\Gamma(j+a)}{\Gamma(j+b)} \sim j^{a-b}$ for j large (see Baillie, 1996, p. 20), we see that for t and $t-j$ large,

$$\kappa_{t,j} \sim \theta \left(t+1\right)^{-\theta} \left(t-j\right)^{-(1-\theta)}.$$

Hence, since $\theta > 0$, as $t \to \infty$,

$$\sum_{j=0}^{t} \kappa_{t,j} \sim \theta t^{-\theta} \sum_{j=1}^{t-1} j^{\theta-1} \to \theta$$

Therefore for t and $t-j$ large, the least-squares weights satisfy:

$$w_{t,j} = \frac{\kappa_{t,j}}{\sum_{i=0}^{t} \kappa_{t,i}} \sim t^{-\theta} \left(t-j\right)^{-(1-\theta)}.$$

2 Proof of Expression (11)

Using expressions (1) and (6), $a_t = \frac{1-g_t}{1-\beta g_t} a_{t-1} + \frac{g_t}{1-\beta g_t} x_t$ such that, for $g_t = \min\left(1, \theta/t\right)$,

$$\begin{aligned} a_t &= a_0 \prod_{i=0}^{t-\lceil\theta\rceil} \frac{1-g_{t-i}}{1-\beta g_{t-i}} + \sum_{j=0}^{t-\lceil\theta\rceil} \left[\frac{g_{t-j}}{1-\beta g_{t-j}} \prod_{i=0}^{j-1} \frac{1-g_{t-i}}{1-\beta g_{t-i}} \right] x_{t-j} \\ &= a_0 \prod_{j=\lceil\theta\rceil}^{t} \frac{1-g_j}{1-\beta g_j} + \sum_{i=\lceil\theta\rceil}^{t} \prod_{j=i+1}^{t} \frac{1-g_j}{1-\beta g_j} \frac{g_i x_i}{1-\beta g_i}, \end{aligned}$$

hence,

$$\begin{aligned} a_t &= a_0 \prod_{j=\lceil\theta\rceil}^{t} \frac{j-\theta}{j-\beta\theta} + \sum_{i=\lceil\theta\rceil}^{t} \prod_{j=i+1}^{t} \frac{j-\theta}{j-\beta\theta} \frac{\theta}{i} \frac{i}{i-\beta\theta} x_i \\ &= \frac{\Gamma\left(t+1-\theta\right)}{\Gamma\left(\lceil\theta\rceil-\theta\right)} \frac{\Gamma\left(\lceil\theta\rceil-\beta\theta\right)}{\Gamma\left(t+1-\beta\theta\right)} a_0 + \theta \frac{\Gamma\left(t+1-\theta\right)}{\Gamma\left(t+1-\beta\theta\right)} \sum_{i=\lceil\theta\rceil}^{t} \frac{\Gamma\left(i-\beta\theta\right)}{\Gamma\left(i+1-\theta\right)} x_i \end{aligned}$$

since $t = \frac{\Gamma(t+1)}{\Gamma(t)}$ so $a\left(a+1\right)....b = \frac{\Gamma(b+1)}{\Gamma(a)}$. Substituting for $y_{t+1}^e = a_t$ into (1) yields

$$y_t = \beta \frac{\Gamma\left(t+1-\theta\right)}{\Gamma\left(\lceil\theta\rceil-\theta\right)} \frac{\Gamma\left(\lceil\theta\rceil-\beta\theta\right)}{\Gamma\left(t+1-\beta\theta\right)} a_0 + x_t + \beta\theta \frac{\Gamma\left(t+1-\theta\right)}{\Gamma\left(t+1-\beta\theta\right)} \sum_{i=\lceil\theta\rceil}^{t} \frac{\Gamma\left(i-\beta\theta\right)}{\Gamma\left(i+1-\theta\right)} x_i.$$

Assume $a_0 = 0$ without loss of generality. Changing the index of x_i and leading this expression j periods yields

$$y_{t+j} = x_{t+j} + \beta\theta \frac{\Gamma\left(t+j+1-\theta\right)}{\Gamma\left(t+j+1-\beta\theta\right)} \sum_{i=\lceil\theta\rceil}^{t+j} \frac{\Gamma\left(i-\beta\theta\right)}{\Gamma\left(i+1-\theta\right)} x_i. \qquad \text{(S-2)}$$

Hence, for $j > 1$ and $t \geq \lceil \theta \rceil$,

$$\frac{\partial y_{t+j}}{\partial x_t} = \beta\theta \frac{\Gamma(t+j+1-\theta)}{\Gamma(t+j+1-\beta\theta)} \frac{\Gamma(t-\beta\theta)}{\Gamma(t+1-\theta)},$$

which yields equation (11) using Stirling's formula (see Baillie, 1996, p. 20).

3 Degree of Long Memory of y_t^2

The next proposition shows that, under Gaussianity, when the degree of long memory of y_t in a model with learning is sufficiently high ($d \geq 1/4$), the long memory of y_t^2, d_2, corresponds to the degree of long memory of a squared ARFIMA(p,d,q), see Haldrup and Kruse (2014).

Proposition 1 *Let the process y_t satisfy the conditions of Theorem 2 with $x_t \overset{iid}{\sim} \mathsf{N}(0, \sigma_x^2)$, $a_0 = 0$ and $(1-\beta)\theta \in (0,1]$. If y_t has long memory of degree $d > 0$, then y_t^2 has long memory of degree*

$$d_2 = \max\left[2\left(d - \frac{1}{2}\right) + \frac{1}{2}, 0\right] = \max\left[2d - \frac{1}{2}, 0\right].$$

Proof The proof proceeds in the following steps. Decompose

$$\sum_{t=1}^{T} y_t^2 = \beta^2 \sum_{t=1}^{T} a_t^2 + 2\beta \sum_{t=1}^{T} a_t x_t + \sum_{t=1}^{T} x_t^2.$$

Because of the Cauchy-Schwarz inequality, we only look at the magnitudes of $\mathsf{var}\left(\sum_{t=1}^{T} x_t^2\right)$, $\mathsf{var}\left(\sum_{t=1}^{T} a_t^2\right)$ and $\mathsf{var}\left(\sum_{t=1}^{T} a_t x_t\right)$ as the covariances will be of no greater magnitude than the variances with highest magnitudes.

First, since $x_t \overset{iid}{\sim} \mathsf{N}(0, \sigma_x^2)$,

$$\begin{aligned}
\mathsf{Var}\left(\sum_{t=1}^{T} x_t^2\right) &= \mathsf{E}\left[\sum_{t=1}^{T} x_t^4 + 2\sum_{t=1}^{T-1}\sum_{j=t+1}^{T} x_t^2 x_j^2\right] - \mathsf{E}\left[\sum_{t=1}^{T} x_t^2\right]^2 \\
&= 3T\sigma_x^4 + 2\sigma_x^4\left[(T-1)T - \frac{(T-1)T}{2}\right] - T^2\sigma_x^4 \\
&= 2T\sigma_x^4. \qquad \text{(S-3)}
\end{aligned}$$

Next, we prove below the following two expressions:

$$\begin{cases}
\mathsf{Var}\left(\sum_{t=1}^{T} a_t^2\right) \asymp T^{2(1-2(1-\beta)\theta)}, & \text{if } (1-\beta)\theta < \frac{1}{2}, \\
\mathsf{Var}\left(\sum_{t=1}^{T} a_t^2\right) = O\left(\log^4 T\right), & \text{if } (1-\beta)\theta = \frac{1}{2}, \\
\mathsf{Var}\left(\sum_{t=1}^{T} a_t^2\right) \asymp T^{2((1-\beta)\theta-1/2)}, & \text{if } (1-\beta)\theta \in \left(\frac{1}{2}, 1\right],
\end{cases} \qquad \text{(S-4)}$$

and, for $(1-\beta)\,\theta \in (0,1]$,

$$\mathsf{Var}\left(\sum_{t=1}^{T} a_t x_t\right) \asymp \log^2 T. \tag{S-5}$$

From (S-3), (S-4), (S-5) and the Cauchy-Schwarz inequality, it follows that the order of magnitude of $\mathsf{Var}\left(\sum_{t=1}^{T} y_t^2\right)$ is the same as that of $\mathsf{Var}\left(\sum_{t=1}^{T} a_t^2\right) + \mathsf{Var}\left(\sum_{t=1}^{T} x_t^2\right)$, so

$$\mathsf{sd}\left(T^{-1/2}\sum_{t=1}^{T} y_t^2\right) \asymp \begin{cases} T^{\frac{1}{2}-2(1-\beta)\theta}, & \text{if } (1-\beta)\,\theta < \frac{1}{4}, \\ 1, & \text{if } (1-\beta)\,\theta \geq \frac{1}{4}. \end{cases} \tag{S-6}$$

Hence, the degree of long memory of y_t^2 is

$$d_2 = \max\left[\frac{1}{2} - 2\,(1-\beta)\,\theta, 0\right] = \max\left(2d - \frac{1}{2}, 0\right);$$

Proof of (S-4) To prove (S-4), we show that

$$\mathsf{E}\left[\left(\sum_{t=1}^{T} a_t^2\right)^2\right] \asymp \begin{cases} T^{2(1-2(1-\beta)\theta)}, & \text{if } \theta\,(1-\beta) < \frac{1}{2}, \\ \log^4 T, & \text{if } \theta\,(1-\beta) = \frac{1}{2}, \\ T^{2((1-\beta)\theta-1/2)}, & \text{if } \theta\,(1-\beta) \in \left(\frac{1}{2}, 1\right], \end{cases} \tag{S-7}$$

and that

$$\left(\mathsf{E}\left[\sum_{t=1}^{T} a_t^2\right]\right)^2 \asymp \begin{cases} T^{2(1-2(1-\beta)\theta)}, & \text{if } (1-\beta)\,\theta < \frac{1}{2}, \\ \log^4 T, & \text{if } (1-\beta)\,\theta = \frac{1}{2}, \\ \log^2 T, & \text{if } (1-\beta)\,\theta > \frac{1}{2}. \end{cases} \tag{S-8}$$

where, when $(1-\beta)\,\theta < 1/2$, there is no reduction in magnitude when computing $\mathsf{E}\left[\left(\sum_{t=1}^{T} a_t^2\right)^2\right] - \left(\mathsf{E}\left[\sum_{t=1}^{T} a_t^2\right]\right)^2$ so $\mathsf{Var}\left(\sum_{t=1}^{T} a_t^2\right)^2 \asymp T^{2(1-2(1-\beta)\theta)}$.

We start with the proof of expression (S-7). Assuming that $a_0 = 0$, we write

$$a_t = h_t \sum_{i=1}^{t} \omega_i x_i,$$

where $\omega_j \sim \frac{\Gamma(j-(1-(1-\beta))\theta)}{\Gamma(j+1-\theta)}$ and $h_j \sim \theta \frac{\Gamma(j+1-\theta)}{\Gamma(j+1-(1-(1-\beta))\theta)}$. Stirling's formula implies that as $j \to \infty$

$$\omega_j \sim j^{-(1-(1-\beta)\theta)}, \tag{S-9a}$$

$$h_j \sim \theta j^{-(1-\beta)\theta}. \tag{S-9b}$$

which we will use several times in the proof.

First, consider

$$\mathsf{E}\left[\left(\sum_{t=1}^{T} a_t^2\right)^2\right] = \sum_{t=1}^{T} \mathsf{E}\left[a_t^4\right] + 2\sum_{t=1}^{T-1}\sum_{j=t+1}^{T} \mathsf{E}\left[a_t^2 a_j^2\right] \tag{S-10}$$

where

$$a_t^2 = \left(h_t \sum_{i=1}^{t} \omega_i x_i\right)^2 = h_t^2 \sum_{i=1}^{t} \omega_i^2 x_i^2 + 2h_t^2 \sum_{i=1}^{t-1}\sum_{j=i+1}^{t} \omega_i \omega_j x_i x_j.$$

Hence, for $j \geq t$, the product $a_t^2 a_j^2$ writes

$$\begin{aligned}
a_t^2 a_j^2 &= h_t^2 h_j^2 \left[\sum_{i=1}^{t} \omega_i^2 x_i^2 + 2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t} \omega_i \omega_k x_i x_k\right]\left[\sum_{i=1}^{j} \omega_i^2 x_i^2 + 2\sum_{i=1}^{j-1}\sum_{k=i+1}^{j} \omega_i \omega_k x_i x_k\right] \\
&= h_t^2 h_j^2 \left[\sum_{i=1}^{t} \omega_i^2 x_i^2 + 2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t} \omega_i \omega_k x_i x_k\right] \\
&\times \left[\sum_{i=1}^{t} \omega_i^2 x_i^2 + \sum_{i=t+1}^{j} \omega_i^2 x_i^2 + 2\sum_{i=1}^{t-1}\left(\sum_{k=i+1}^{t} \omega_i \omega_k x_i x_k + \sum_{k=t+1}^{j} \omega_i \omega_k x_i x_k\right)\right. \\
&\left. + \sum_{i=t}^{j-1}\sum_{k=i+1}^{j} \omega_i \omega_k x_i x_k\right].
\end{aligned}$$

Collecting terms,

$$\begin{aligned}
a_t^2 a_j^2 &= h_t^2 h_j^2 \left[\sum_{i=1}^{t} \omega_i^2 x_i^2 + 2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t} \omega_i \omega_k x_i x_k\right]^2 \\
&+ h_t^2 h_j^2 \left[\sum_{i=1}^{t} \omega_i^2 x_i^2 + 2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t} \omega_i \omega_k x_i x_k\right] \\
&\times \left[\sum_{i=t+1}^{j} \omega_i^2 x_i^2 + 2\sum_{i=1}^{t-1}\sum_{k=t+1}^{j} \omega_i \omega_k x_i x_k + \sum_{i=t}^{j-1}\sum_{k=i+1}^{j} \omega_i \omega_k x_i x_k\right].
\end{aligned}$$

The expectation of the product $a_t^2 a_j^2$ can therefore be decomposed in two simpler terms:

$$\begin{aligned}
\mathsf{E}\left[a_t^2 a_j^2\right] &= h_t^2 h_j^2 \mathsf{E}\left[\sum_{i=1}^{t} \omega_i^2 x_i^2 + 2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t} \omega_i \omega_k x_i x_k\right]^2 \\
&+ h_t^2 h_j^2 \mathsf{E}\left[\left(\sum_{i=1}^{t} \omega_i^2 x_i^2\right)\left(\sum_{i=t+1}^{j} \omega_i^2 x_i^2\right)\right]
\end{aligned} \tag{S-11}$$

since all the other elements in the previous expressions have zero expectations. We consider in turn the two terms that appear on the right-hand side of expression (S-11). First, progressively expanding the quadratic terms and discarding the elements with zero expectations, we obtain the following simplification:

$$
\begin{aligned}
&\mathsf{E}\left[\sum_{i=1}^{t}\omega_i^2x_i^2+2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i\omega_kx_ix_k\right]^2\\
&=\mathsf{E}\left(\sum_{i=1}^{t}\omega_i^2x_i^2\right)^2+\mathsf{E}\left(2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i\omega_kx_ix_k\right)^2\\
&=\mathsf{E}\left(\sum_{i=1}^{t}\omega_i^4x_i^4+2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i^2\omega_k^2x_i^2x_k^2\right)\\
&+4\mathsf{E}\left(\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i^2\omega_k^2x_i^2x_k^2\right)\\
&=3\sigma_x^4\left[\sum_{i=1}^{t}\omega_i^4+2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i^2\omega_k^2\right].
\end{aligned}
$$

The second term in expression (S-11) satisfies

$$
\mathsf{E}\left[\left(\sum_{i=1}^{t}\omega_i^2x_i^2\right)\left(\sum_{i=t+1}^{j}\omega_i^2x_i^2\right)\right]=\left[\sum_{i=1}^{t}\omega_i^2\sum_{s=t+1}^{j}\omega_s^2\right]\sigma_x^4.
$$

Collecting the previous results, for $j \geq t$,

$$
\begin{aligned}
\mathsf{E}\left[a_t^2a_j^2\right]=3\sigma_x^4h_t^2h_j^2&\left[\sum_{i=1}^{t}\omega_i^4+2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i^2\omega_k^2\right]\\
&+h_t^2h_j^2\left[\sum_{i=1}^{t}\omega_i^2\sum_{s=t+1}^{j}\omega_s^2\right]\sigma_x^4.
\end{aligned}
$$

Expression (S-10) therefore writes:

$$
\begin{aligned}
\mathsf{E}\left(\sum_{t=1}^{T}a_t^2\right)^2&=3\sigma_x^4\sum_{t=1}^{T}h_t^4\left[\sum_{i=1}^{t}\omega_i^4+2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i^2\omega_k^2\right] &\text{(S-12)}\\
&+6\sigma_x^4\sum_{t=1}^{T-1}\sum_{j=t+1}^{T}h_t^2h_j^2\left[\sum_{i=1}^{t}\omega_i^4+2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i^2\omega_k^2\right] &\text{(S-13)}\\
&+2\sigma_x^4\sum_{t=1}^{T-1}\sum_{j=t+1}^{T}h_t^2h_j^2\left(\sum_{i=1}^{t}\omega_i^2\right)\left(\sum_{s=t+1}^{j}\omega_s^2\right) &\text{(S-14)}
\end{aligned}
$$

We consider, in turn, the three cases where (a) $(1-\beta)\,\theta \in (0,1/2)$, (b) $(1-\beta)\,\theta = 1/2$, and (c) $(1-\beta)\,\theta \in (1/2,1]$. We study the asymptotic behaviors implied by those of h_j and ω_j given in expression (S-9).

First, case (a), when $(1-\beta)\,\theta < 1/2$ so $2\left((1-\beta)\,\theta - 1\right) < -1$ then as $t \to \infty$

$$\sum_{i=1}^{t} \omega_i^4 \sim \sum_{i=1}^{t} i^{4((1-\beta)\theta-1)} \asymp 1,$$

and

$$\begin{aligned}\sum_{i=1}^{t-1} \sum_{k=i+1}^{t} \omega_i^2 \omega_k^2 &\sim \sum_{i=1}^{t-1} \sum_{k=i+1}^{t} k^{2((1-\beta)\theta-1)} i^{2((1-\beta)\theta-1)} \\ &= \sum_{k=2}^{t} \sum_{i=1}^{k-1} k^{2((1-\beta)\theta-1)} i^{2((1-\beta)\theta-1)} \\ &\asymp 1.\end{aligned}$$

Hence expression (S-12) satisfies, as $T \to \infty$,

$$\begin{aligned}&3\sigma_x^4 \sum_{t=1}^{T} h_t^4 \left[\sum_{i=1}^{t} \omega_i^4 + 2\sum_{i=1}^{t-1} \sum_{k=i+1}^{t} \omega_i^2 \omega_k^2\right] \asymp \sum_{t=1}^{T} h_t^4 \\ &\asymp \sum_{t=1}^{T} t^{-4(1-\beta)\theta} \\ &\asymp \begin{cases} T^{1-4(1-\beta)\theta}, & \text{if } (1-\beta)\,\theta < 1/4; \\ \log T, & \text{if } (1-\beta)\,\theta = 1/4; \\ 1, & \text{if } (1-\beta)\,\theta > 1/4. \end{cases}\end{aligned}$$

Consider now expression (S-13), as $T \to \infty$. Since $2\,(1-\beta)\,\theta \in (0,1)$,

$$\begin{aligned}&6\sigma_x^4 \sum_{t=1}^{T-1} \sum_{j=t+1}^{T} h_t^2 h_j^2 \left[\sum_{i=1}^{t} \omega_i^4 + 2\sum_{i=1}^{t-1} \sum_{k=i+1}^{t} \omega_i^2 \omega_k^2\right] \\ &\asymp \sum_{t=1}^{T-1} \sum_{j=t+1}^{T} h_t^2 h_j^2 \asymp \sum_{t=1}^{T-1} \sum_{j=t+1}^{T} t^{-2(1-\beta)\theta} j^{-2(1-\beta)\theta} \\ &\asymp T^{2(1-2(1-\beta)\theta)}\end{aligned}$$

Finally, expression (S-14) satisfies

$$2\sigma_x^4 \sum_{t=1}^{T-1} \sum_{j=t+1}^{T} h_t^2 h_j^2 \left(\sum_{i=1}^{t} \omega_i^2\right) \left(\sum_{s=t+1}^{j} \omega_s^2\right) = O\left(T^{2(1-2(1-\beta)\theta)}\right)$$

since $\left(\sum_{i=1}^{t}\omega_i^2\right)\left(\sum_{s=t+1}^{j}\omega_s^2\right) = O(1)$ as $t\to\infty$. It follows, as $2(1-2(1-\beta)\theta) > 1 - 4(1-\beta)\theta$, that

$$\mathsf{E}\left[\left(\sum_{t=1}^{T} a_t^2\right)^2\right] \asymp T^{2(1-2(1-\beta)\theta)}.$$

Now, consider the case (b) where $(1-\beta)\theta = 1/2$, then $\sum_{i=1}^{t}\omega_i^4 \asymp 1$ and, as $t\to\infty$,

$$\begin{aligned}\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i^2\omega_k^2 &\sim \sum_{i=1}^{t-1}\sum_{k=i+1}^{t} k^{-1}i^{-1}\\ &= \sum_{k=2}^{t}\frac{\log k}{k} \asymp \log^2 t.\end{aligned}$$

Hence expression (S-12) satisfies, as $T\to\infty$,

$$3\sigma_x^4\sum_{t=1}^{T}h_t^4\left[\sum_{i=1}^{t}\omega_i^4 + 2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i^2\omega_k^2\right] \asymp \sum_{t=1}^{T} t^{-2}\log^2 t \asymp T^{-1}\log^2 T.$$

Now, regarding expression (S-13) since $2(1-\beta)\theta = 1$,

$$\begin{aligned}&6\sigma_x^4\sum_{t=1}^{T-1}\sum_{j=t+1}^{T}h_t^2h_j^2\left[\sum_{i=1}^{t}\omega_i^4 + 2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t}\omega_i^2\omega_k^2\right]\\ &\asymp \sum_{t=1}^{T-1}\frac{\log^2 t}{t}\sum_{j=t+1}^{T} j^{-1}\\ &\asymp \log^4 T\end{aligned}$$

and, for expression (S-14),

$$\begin{aligned}&2\sigma_x^4\sum_{t=1}^{T-1}\sum_{j=t+1}^{T}h_t^2h_j^2\left(\sum_{i=1}^{t}\omega_i^2\right)\left(\sum_{s=t+1}^{j}\omega_s^2\right)\\ &\sim 2\sigma_x^4\theta^2\sum_{t=1}^{T-1}\sum_{j=t+1}^{T}\left(\frac{\log t}{t}\right)\left(\frac{\log j - \log t}{j}\right)\\ &\asymp \sum_{t=1}^{T-1}\left(\frac{\log t}{t}\right)\sum_{j=t+1}^{T}\left(\frac{\log j - \log t}{j}\right)\\ &\asymp \log^4 T.\end{aligned}$$

hence

$$\mathsf{E}\left(\sum_{t=1}^{T} a_t^2\right)^2 \asymp \log^4 T.$$

Consider now the case (c) where $(1-\beta)\theta \in (1/2, 1]$, then as $t \to \infty$,

$$\sum_{i=1}^{t} \omega_i^4 \sim \sum_{i=1}^{t} i^{4((1-\beta)\theta-1)} \asymp \begin{cases} 1, & \text{if } (1-\beta)\theta \in (1/2, 3/4); \\ \log t, & \text{if } (1-\beta)\theta = 3/4; \\ t^{4((1-\beta)\theta-3/4)}, & \text{if } (1-\beta)\theta \in (3/4, 1). \end{cases},$$

and

$$\begin{aligned} \sum_{i=1}^{t-1} \sum_{k=i+1}^{t} \omega_i^2 \omega_k^2 &\sim \sum_{i=2}^{t} \sum_{k=1}^{k-1} k^{2((1-\beta)\theta-1)} i^{2((1-\beta)\theta-1)} \\ &\asymp t^{4((1-\beta)\theta-1/2)}. \end{aligned}$$

Hence, for expression (S-12),

$$\begin{aligned} 3\sigma_x^4 \sum_{t=1}^{T} h_t^4 \left[\sum_{i=1}^{t} \omega_i^4 + 2 \sum_{i=1}^{t-1} \sum_{k=i+1}^{t} \omega_i^2 \omega_k^2 \right] &\asymp \sum_{t=1}^{T} t^{-4(1-\beta)\theta} t^{4(1-\beta)\theta-2} \\ &\asymp 1 \end{aligned}$$

Now, for expression (S-13), with $2(1-\beta)\theta > 1$ and as $T \to \infty$,

$$\begin{aligned} &6\sigma_x^4 \sum_{t=1}^{T-1} \sum_{j=t+1}^{T} h_t^2 h_j^2 \left[\sum_{i=1}^{t} \omega_i^4 + 2 \sum_{i=1}^{t-1} \sum_{k=i+1}^{t} \omega_i^2 \omega_k^2 \right] \\ &\asymp \sum_{t=1}^{T-1} \sum_{j=t+1}^{T} t^{2((1-\beta)\theta-1)} j^{-2(1-\beta)\theta} \\ &\asymp \sum_{t=2}^{T} \sum_{j=1}^{t-1} t^{2((1-\beta)\theta-1)} j^{-2(1-\beta)\theta} \asymp \sum_{t=2}^{T} t^{2((1-\beta)\theta-1)} \\ &\asymp T^{2((1-\beta)\theta-1/2)}. \end{aligned}$$

Finally, expression (S-14) yields

$$\begin{aligned} &2\sigma_x^4 \sum_{t=1}^{T-1} \sum_{j=t+1}^{T} h_t^2 h_j^2 \left(\sum_{i=1}^{t} \omega_i^2 \right) \left(\sum_{s=t+1}^{j} \omega_s^2 \right) \\ &\sim 2\sigma_x^4 \sum_{t=1}^{T-1} \sum_{j=t+1}^{T} h_t^2 h_j^2 \left(\sum_{i=1}^{t} i^{-2(1-(1-\beta)\theta)} \right) \left(\sum_{s=t+1}^{j} s^{-2(1-(1-\beta)\theta)} \right) \\ &\sim \frac{2\sigma_x^4 \theta^4}{4\left((1-\beta)\theta - 1/2\right)^2} \sum_{t=2}^{T} \sum_{j=1}^{t-1} t^{-1} \left(j^{-1} - j^{-2(1-\beta)\theta} t^{2(1-\beta)\theta-1} \right) \\ &\asymp T^{2(1-\beta)\theta-1}. \end{aligned}$$

It follows that, since the exact $T^{2(1-\beta)\theta-1}$ magnitude does not cancel out when adding (S-13) and (S-14), that

$$\mathsf{E}\left(\sum_{t=1}^{T} a_t^2\right)^2 \asymp T^{2(1-\beta)\theta-1}. \tag{S-15}$$

This completes the proof of expression (S-7).

We now turn to the proof of expression (S-8). Consider

$$\mathsf{E}\left[\sum_{t=1}^{T} a_t^2\right] = \sigma_x^2 \sum_{t=1}^{T}\left(h_t^2 \sum_{i=1}^{t} \omega_i^2\right)$$

where

$$\omega_j^2 \sim j^{-2(1-(1-\beta)\theta)},$$
$$h_j^2 \sim \theta j^{-2(1-\beta)\theta}.$$

It follows immediately that

$$\mathsf{E}\left[\sum_{t=1}^{T} a_t^2\right] \asymp \begin{cases} T^{1-2(1-\beta)\theta}, & \text{if } \theta\,(1-\beta) < \frac{1}{2}, \\ \log^2 T, & \text{if } \theta\,(1-\beta) = \frac{1}{2}, \\ \log T, & \text{if } \theta\,(1-\beta) > \frac{1}{2}. \end{cases}$$

When $(1-\beta)\,\theta < 1/2$, there is no reduction in magnitude when computing $\mathsf{E}\left[\left(\sum_{t=1}^{T} a_t^2\right)^2\right] - \left(\mathsf{E}\left[\sum_{t=1}^{T} a_t^2\right]\right)^2$, indeed the magnitude of $\mathsf{E}\left[\left(\sum_{t=1}^{T} a_t^2\right)^2\right]$ is driven by expression (S-13):

$$6\sigma_x^4 \sum_{t=1}^{T-1}\sum_{j=t+1}^{T} h_t^2 h_j^2 \left[\sum_{i=1}^{t} \omega_i^4 + 2\sum_{i=1}^{t-1}\sum_{k=i+1}^{t} \omega_i^2\omega_k^2\right]$$

which differs from the driving term in the magnitude of $\left(\mathsf{E}\left[\sum_{t=1}^{T} a_t^2\right]\right)^2$, namely $\sigma_x^4\left[\sum_{t=1}^{T} h_t^2 \sum_{i=1}^{t} \omega_i^2\right]^2$ owing to the factor $6\sigma_x^4$ in front of the expression.

Proof of (S-5) We now turn to $\mathsf{Var}\left(\sum_{t=1}^{T} a_t x_t\right)$, noticing that $\mathsf{E}\left[\sum_{t=1}^{T} a_t x_t\right] = \sigma_x^2 \sum_{t=1}^{T} h_t \omega_t \asymp \log T$.

Now, since $a_t = h_t \sum_{i=1}^t \omega_i x_i$,

$$\left(\sum_{t=1}^T a_t x_t\right)^2 = \sum_{t=1}^T a_t^2 x_t^2 + 2\sum_{t=2}^T \sum_{k=1}^{t-1} a_t x_t a_k x_k$$
$$= \sum_{t=1}^T h_t^2 \left[\sum_{i=1}^t \omega_i^2 x_i^2 + 2\sum_{i=1}^{t-1} \sum_{j=i+1}^{t} \omega_i \omega_j x_i x_j\right] x_t^2$$
$$+ 2\sum_{t=2}^T \sum_{k=1}^{t-1} \sum_{i=1}^{t} \sum_{j=1}^{k} h_t h_k \omega_i \omega_j x_i x_t x_j x_k$$

Hence, taking expectations, the expression simplifies to

$$\mathsf{E}\left[\left(\sum_{t=1}^T a_t x_t\right)^2\right] = \sigma_x^4 \sum_{t=1}^T h_t^2 \left[3\omega_t^2 + \sum_{i=1}^{t-1} \omega_i^2\right] + 2\sigma_x^4 \sum_{t=2}^T \sum_{k=1}^{t} h_t h_k \omega_t \omega_k.$$

Now, using the same asymptotic expressions for h_j and ω_j as before, the previous expression becomes

$$\mathsf{E}\left[\left(\sum_{t=1}^T a_t x_t\right)^2\right] \sim \sigma_x^4 \theta^2 \sum_{t=1}^T t^{-2(1-\beta)\theta} \left[2t^{-2(1-(1-\beta)\theta)} + \sum_{i=1}^t i^{-2(1-(1-\beta)\theta)}\right] + 2\theta^2 \sigma_x^4 \sum_{t=2}^T \sum_{k=1}^t t^{-1} k^{-1}$$
$$\asymp \sigma_x^4 \theta^2 \sum_{t=1}^T \left[2t^{-2} + t^{-2(1-\beta)\theta} \sum_{i=1}^t i^{-2(1-(1-\beta)\theta)}\right] + \log^2 T.$$

We use

$$\sum_{i=1}^t i^{-2(1-(1-\beta)\theta)} \asymp \begin{cases} 1, & \text{if } \theta(1-\beta) < \frac{1}{2}, \\ \log t, & \text{if } \theta(1-\beta) = \frac{1}{2}, \\ t^{(1-\beta)\theta - 1}, & \text{if } \theta(1-\beta) > \frac{1}{2}. \end{cases}$$

Hence, for all $(1-\beta)\theta \in (0,1]$

$$\mathsf{E}\left[\left(\sum_{t=1}^T a_t x_t\right)^2\right] \asymp \log^2 T,$$

and the result follows.

4 Additional Simulations

We present two complementary figures that complement the results of Theorem 2 in the paper. Figures S.1 and S.2 report, respectively, the log of $\mathsf{sd}\left(T^{-1/2}\sum_{t=1}^T y_t\right)$ and the growth

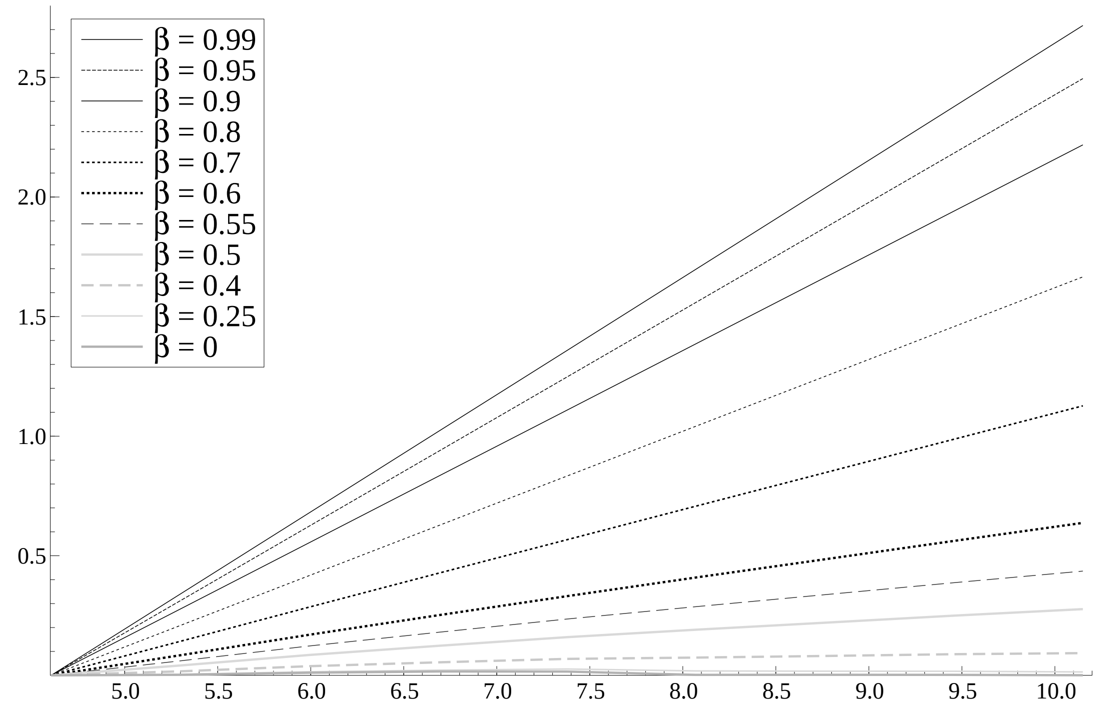

Figure S.1: Magnitude of the log of the Monte Carlo standard deviation of the sample mean, $\log \widehat{\mathsf{sd}} \left(T^{-1/2} \sum y_t \right)$, as a function of the log sample size when agents learn using RLS.

rate of $\mathsf{sd} \left(T^{-1/2} \sum_{t=1}^{T} y_t \right) / \log T$ under RLS learning as a function of $\log T$. In both figures, the horizontal axes is the log of the sample size. The sample sizes range from 200 to 50,000 for which we produce 10,000 Monte Carlo replications. The data generating process is the same as in the Simulations section (Section 4) of the paper, with $\sigma_0^2 = \infty$. The figures illustrate that, as the sample size increases, $\mathsf{sd} \left(T^{-1/2} \sum_{t=1}^{T} y_t \right)$ behaves as the Theorem 2 implies.

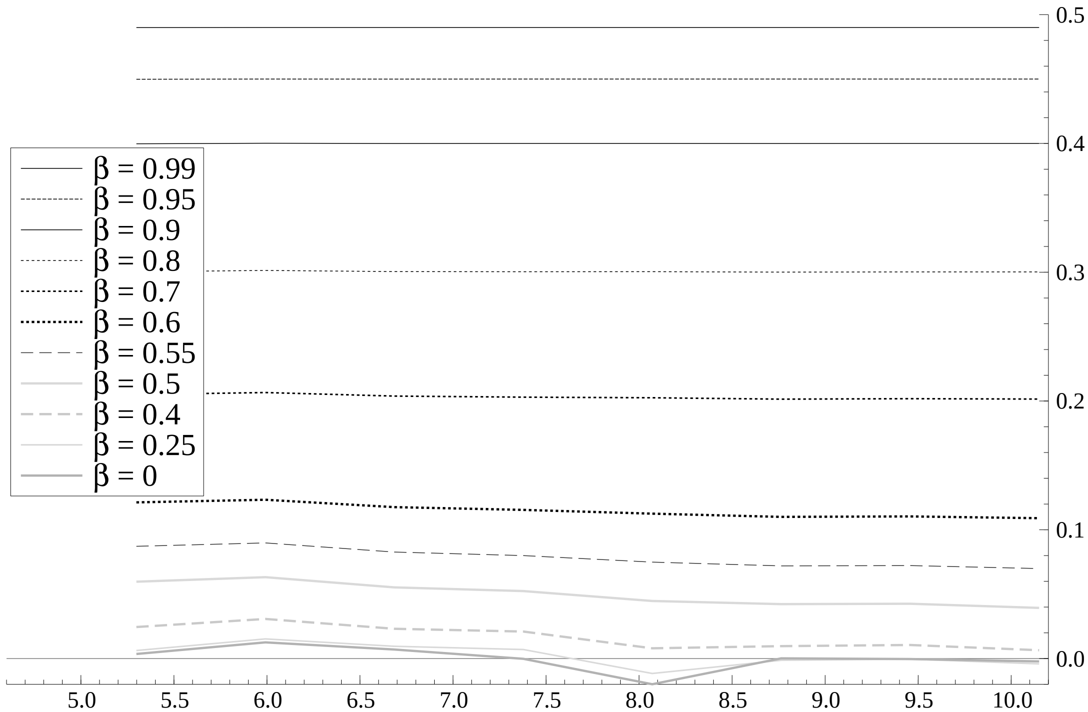

Figure S.2: Growth of the ratio $\log \widehat{\mathsf{sd}} \left(T^{-1/2} \sum y_t\right) / \log T$ as a function of the log sample size T when agents learn using RLS.